Journal of Multidisciplinary Engineering Science Studies (JMESS)
ISSN: 2458-925X
Vol. 6 Issue 11, November - 2020

Artificial Neural Network and Time Series Modeling of predicted future production with lentil production in Turkey and Analysis

Şenol Çelik
Bingol University Faculty of Agriculture,
Department of Biometry and Genetics,
Bingol, Turkey
senolcelik@bingol.edu.tr

***Abstract*— The goal of this study is to display that production design may be performed using artificial neural networks (ANN) and time series analysis (ARIMA) in the establishment of lentils production amount model and in forecasting in Turkey by years.**

In the development of ANN and ARIMA, parameter of years was used as an independent parameter and production amount was used as a dependent parameter. The suitable of the model developed was determined using Square Mean Squared Error (RMSE). The results show that lentils production will have a fluctuating course in production from 2020 to 2025.

It was observed that better results were obtained from ANN method than time series analysis in lentils production prediction in Turkey.

Keywords*— *ANN; time series analysis; production; lentil

I. Introduction (*Heading 1*)

Legumes are one of the important food and feed sources in closing the protein gap in animal nutrition with its protein value in both human nutrition and straw. Especially in the world on the agenda in recent years, "Whole-Food Plant-Based Nutrition (WFPB)" In Turkey, "Holistic Plant Based Nutrition (4B)" as a known vegan diet is becoming increasingly popular. However, the implementation of quarantine almost all over the world due to the pandemic has directed consumers to storable products. With the onset of the pandemic effect in the world, there has been an intense increase in demand for legumes [1]. According to 2018 FAO data, 6.33 million tons of production was realized on 6.1 million hectares [2].

As can be seen from Table 1, Canada ranks first in the world lentils rankings. It is followed by India in the 2nd place, United States of America in the third place, Turkey in the fourth place and Australia in the fifth place. Canada covers 33.03% of the world production of lentils and the percentage of the production made by the top 10 countries with the highest production is 90.76%.

Table 1. Countries with the highest lentils production in the world

Order	Countries	Production amount (tonnes)	Rate within the world (%)
1	Canada	2 092 136	33.03
2	India	1 620 000	25.58
3	United States of America	381 380	6.02
4	Turkey	353 000	5.57
5	Australia	255 185	4.03
6	Kazakhstan	253 552	4
7	Nepal	249 491	3.94
8	Russian Federation	194 726	3.07
9	Bangladesh	176 633	2.79
10	China	172 173	2.72
	Top 10 countries total	5 748 276	90.76
	Total (World)	**6 333 352**	

There are researchers present on the production model and prediction in agriculture using Artificial Neural Networks (ANN) and Time Series Methods.

The modeling of banana production [3] and forage plant production quantity and prediction were made by ARIMA and exponential smoothing methods [4]. Production volume of nuts with ARIMA models [5], potato [6], red meat [7] and peanut production model [8], orange with artificial neural networks [9], fodder beet [10] and tobacco production [11] modeling and prediction were investigated. Mandarin production [12] using time series and artificial neural networks were modeled. Regarding lentils production, [13, 14] applied time series analysis.

The target of this research is to accomplish the modeling and future prediction for the amount of lentils production in Turkey through the ANN and ARIMA models.

II. MATERIAL AND METHOD

A. *Material*

The material of the study is 1960-2019 lentils production amount values ensured from the www.tuik.gov.tr web address of Turkish Statistical Institute [15]. The dependent variable was dry beans production figures whereas the independent variable was year series. These variables were selected in order to be able to make reasonable estimations with the models to be performed using ANN and time series analysis methods.

B. *Method*

Artificial neural networks (ANN)

Artificial neural networks (ANN) model in is a multi-layer perceptron, a model comprising an input layer, one or more hidden layers, and an output layer. The input layer consists of a set of inputs, each of which represents a direction of the data being modeled. The exact configuration of inputs is an area of investigation in this study, as it is searched to obtain a model with the highest degree of accuracy possible while minimizing the number of inputs needed, simplifying the model.

A neural network utilizes hidden neurons to allow it to model nonlinear functions. The hidden neurons are arranged into one or more hidden layers, the configuration of which depends on the problem at hand, and itself can be cast as an optimization problem. The output layer comprises the neuron(s) that produce the actual output of the model. The output of a neuron is described as following

$$y_k = a\left(\sum_{i=1}^{D} x_i w_{ij} + w_{j0}\right)$$

where a is the activation function and D is model inputs [16].

The utilized activation function in configuration of ANNs in the study is Hyperbolic tangent sigmoid function.

$$f = \frac{2}{1+e^{-net_j}} - 1$$

Purlin function generates outputs in the range of -∞ to +∞, logsig function produces outputs in the range of 0 to 1, and tansig function produces outputs in the range of -1 to +1 [17].

In order to train the ANN, production values parameter were utilized for training and testing. During the training process, the weights were adjusted to make the actual outputs close to the target (measured) output of the network, by the L-M algorithm in accordance with the root mean squared error (RMSE). To evaluate the precision of the predicted discharge volume, Square Mean Square Error (RMSE) [18] were used:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(\hat{I_i} - I_i)^2}{N}}$$

where $\hat{I_i}$ is the estimated discharge for sample i, I_i is the discharge volume obtained from reference data, and N is the number of samples.

Time Series Analysis

A pth-order autoregressive model AR(p) model is denoted as [19].

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + e_t$$

An AR(p) model uses a linear combination of past values of the target to make forecasts.

A qth-order moving average process, expressed MA(q), is characterized by [20].

$$y_t = -\theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q} + e_t$$

ARMA(p,q) model composed of a pth-order autoregressive and qth-order moving average process and it is characterized by [21].

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q}$$

entire proceedings, and not as an independent document. Please do not revise any of the current designations.

III. RESULTS AND DISCUSSION

The artificial neural networks and time series analysis method goodness of fit statistics (RMSE) of lentils production between the years 1960-2019 in Turkey are showed in Table 2.

Table 2. ANN and time series analysis models for lentils production amount

Model	Parameters	Coefficients	RMSE	BIC
ARIMA(1,1,0)	AR(1)	-0.268*	125458	23.618
ARIMA(0,1,1)	MA(1)	0.326*	124349	23.600
ARIMA(1,1,1)	AR(1) and MA(1)	0.115 and 0.428	125338	23.685

*: (p<0.05)

Autocorrelations (ACF) and partial autocorrelation graphs (PACF) applied in time series are given in Figure 1. When Figure 1 is examined, it is displayed that many relation values in the ACF diagram of the series surpass the confidence limit. Namely, there is a trend but not stable. So, the first difference of the series is considered. ACF - and PACF diagrams of the series with the first difference are presented (Figure 2). In Figure 2 all relationship values are within the confidence limits. Goodness of fit statistics such as RMSE and BIC were examined to determine the model grade and the suitable method of the series. The most suitable model is ARIMA(0,1,1). It was compared with the ARIMA model (Table 2).

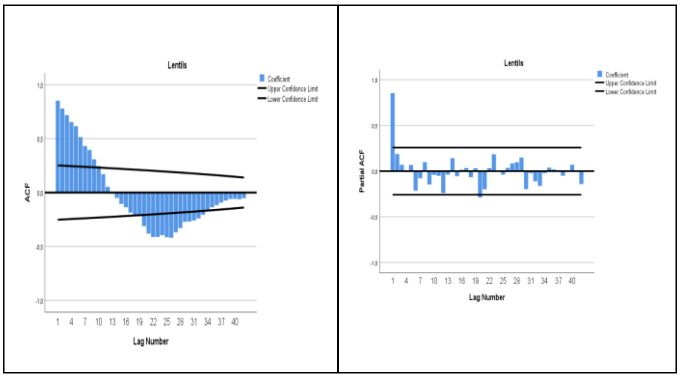

Figure 1. ACF and PACF graph for level series

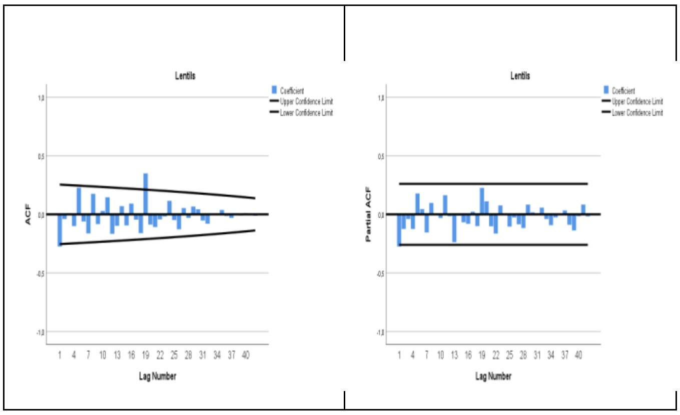

Figure 2. ACF and PACF graph for first difference

When Table 2 is examined, when the time series analysis and artificial neural network methods are compared according to square mean square error (RMSE) values, artificial neural networks (ANN) with minimum RMSE values (RMSE=119175) are the most suitable model. The hyperbolic tangent function was used as activation function when creating a model with the ANN method. The number of neurons in the input layer, the hidden layer and the output layer was determined as 1-12-1 each. 1000 iterations were used for the ANN method in the data series consisting of 60 observations between 1960-2019. The estimated and residual values are presented in Table 3 together with the real values of the ANN method for 2005-2019 period.

Table 3. Observed, estimated and residual values

Years	Actual	Predicted	Residual
2005	570000	546750.12	23249.88
2006	622624	573868.38	48755.62
2007	535181	618075.29	-82894.29
2008	131188	542270.94	-411082.94
2009	302181	162677.54	139503.46
2010	447400	305169.48	142230.52
2011	405952	455657.28	-49705.28
2012	438000	412484.22	25515.78
2013	417000	445943.04	-28943.04
2014	345000	424063.29	-79063.29
2015	360000	348682.22	11317.78
2016	365000	364277.53	722.47
2017	430000	369501.30	60498.70
2018	353000	437633.22	-84633.22
2019	310000	356983.88	-46983.88

The graph of the observed and estimated values obtained with ANN method is showed in Figure 3.

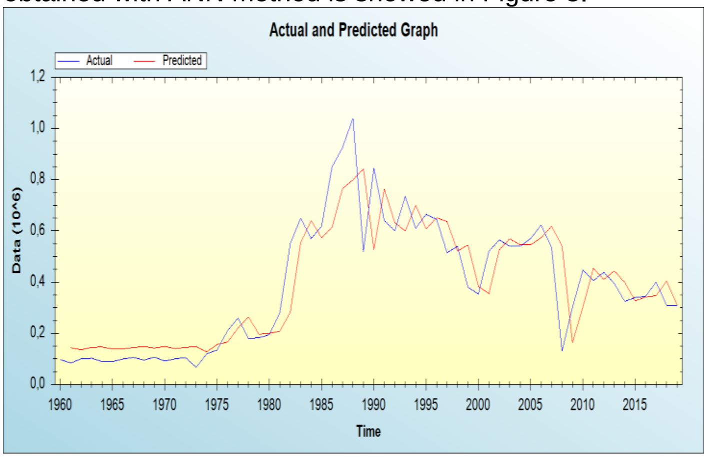

Figure 3. The combined graph of observed and estimated values for lentils production

The possible 2020-2025 values of lentils production forecasted with ANN and ARIMA(0,1,1) are given in Table 4.

Table 4. Lentils production amount forecasting

Years	ARIMA(0,1,1)	ANN
2020	335 954	322971
2021	339 999	325955
2022	344 044	328962
2023	348 089	322002
2024	352 134	335085
2025	356 180	338221

Table 4 shows that lentils production will increase between 2020-2025. The graph showing the actual and predicted values of the lentils production volume is shown in Figure 4.

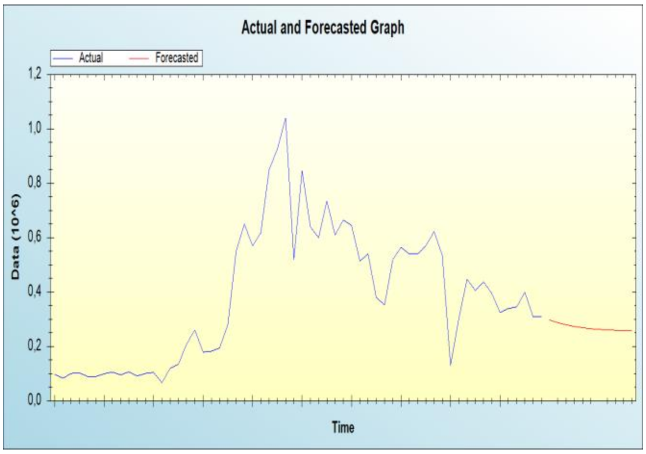

Figure 4. The joint graph of observed and estimated values

As can be showed in Figure 4, lentils production will increase after 2019 and this increase will continue until 2025. In Figure 5, when the joint graph of observed and residual values was observed, residual and actual values were found to be scattered free from each other and randomly. This situation shows that important hypotheses regarding the model are provided.

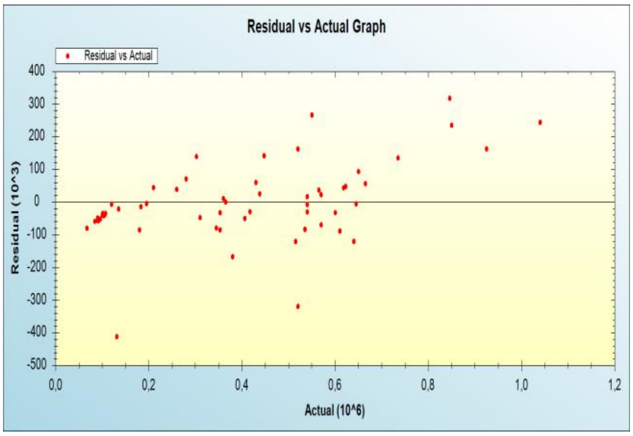

Figure 5. Joint graph of observed and residual values

From our study, we obtained suitable model for forecasting ARIMA (0,1,1) and ANN. After using these models, it can be said that forecasted production will increase to low in future. From [22] study, the best model for forecasting ARIMA (1,1,1), [14] from research, the best model for forecasting ARIMA (0,1,9) and it can be said that forecasted production will increase to in future.

IV. CONCLUSION

Lentils production amount in Turkey was estimated through artificial neural networks and time series analysis such as AR, MA and ARMA in the study. The years (1960-2019) were used as the input variable, as 1 independent variable and lentils production values were used as the output variable. Then, the training, test and verification processes of the network were made and prediction was applied. The results obtained have set out that the ANN model established has given better results than time series analysis methods. Low RMSE statistics in the training, test and verification phases also ground the results.

Taking into account, the estimates of lentils production, it has been noted that the production of 310000 tons in 2019 will increase by 9.1% and reach 338221 tons in 2025. In prediction studies, artificial neural networks can be used as an alternative method to the prediction model in various field.

REFERENCES

[1] Bitkisel Üretim Genel Müdürlüğü Tarım Havzaları Daire Başkanlığı. 2020. Ürün masaları mercimek bülteni, Sayı:6, Nisan 2020.

[2] FAO, Food and Agriculture Organization of the United States. 2018. Production quantity. http://www.fao.org/faostat/en/#data/QC Access: 19 September 2020.

[3]. Eyduran, S. P., Akın, M., Eyduran, E., Çelik, Ş., Ertürk, Y. E., Ercişli, S. 2020. Forecasting Banana Harvest Area and Production in Turkey Using Time Series Analysis. Erwerbs-Obstbau, https://doi.org/10.1007/s10341-020-00490-1

[4]. Ağırbaş, N. C., Çelik, Ş., Sapmaz, K. 2019. Modeling forage crops production using the time series method. Fresenius Environmental Bulletin, 28(11): 7763-7776.

[5]. Çelik, Ş. 2013. Modelling of Production Amount of Nuts Fruit by Using Box-Jenkins Technique. YYU J. Agr. Sci. 23(1): 18-30.

[6]. Çelik, Ş. 2019. Modeling and Estimation of Potato Production in Turkey with Time Series Analysis. International Journal of Trend in Research and Development, 6(5): 111-116.

[7]. Çelik, Ş. 2012. Türkiye de Kırmızı Et Üretiminin Box Jenkins Yöntemiyle Modellenmesi ve Üretim Projeksiyonu. Hayvansal Üretim Dergisi, 53(2):32–39

[8]. Çelik, S., Karadas, K., Eyduran, E. 2017. Forecasting the production of groundnut in Turkey using ARIMA model. The Journal of Animal and Plant Sciences, 27(3): 920-928.

[9]. Çelik, Ş. 2019. Estimation of the orange production in Turkey by means of artificial neural Networks. Global Journal of Engineering Science and Researches, 6(9): 10-16.

[10]. Çelik, Ş. 2020. Estimation modelling of the amount of fodder beet production in Turkey: Comparative Analysis of Artificial Neural Networks and Trend Analysis Methods. Journal of Multidisciplinary Engineering Science Studies (JMESS), 6(7): 3372-3375.

[11]. Çelik, Ş. 2020. Estimation modelling of tobacco production in Turkey: Comparative Analysis of Artificial Neural Networks and Multiplicative Decomposition Methods. International Journal of Trend in Research and Development, 7(4), 154-187.

[12]. Çelik, Ş. 2019. Prediction of Mandarin Production in Turkey through Artificial Neural Networks and Time-

Series Analysis. International Journal of Trend in Research and Development, 6(5): 85-90.

[13]. Berk, A., Şanlı, H., Özcan, İ.A. 2019. Forecasting of lentil production in Turkey importance of cooperatives. 23rd International Turkısh Cooperative Congress September 26-28, 2019 Kiev/Ukraine.

[14]. Farhad Rahman, N. Md., Azizul Baten, Md., Roy, A., Monir Hossain, Md. 2013. Forecasting of Lentil Pulse Production: An Econometric Analysis. Australian Journal of Basic and Applied Sciences, 7(2): 819-824.

[15]. TSI, 2019. Turkey Statistics Institute. Dry pulses, 1988-2019. http://www.tuik.gov.tr/PreTablo.do?alt_id=1001

[16]. Walker, D., Keedwell, E., Savi´c, D., Kellagher, R. 2014. An artificial neural network-based rainfall runoff model for improved drainage network modeling, in: Proc. International Conference on Hydroinformatics (HIC2014).

[17]. Bouabaz, M., Hamami, M. 2008. A cost estimation model for repair bridges based on artificial neural network. American Journal of Applied Science 5(4): 334–339.

[18]. Wang, W., Lu, Y. 2018. Analysis of the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE) in Assessing Rounding Model. ICMEMSCE, IOP Conf. Series: Materials Science and Engineering 324 (2018) 012049, doi:10.1088/1757-899X/324/1/012049.

[19]. Hamilton, J. D. 1994. Time Series Analysis. Princeton University Press Princeton, New Jersey.

[20]. Cooray, T.M.J.A. 2008. Applied Time Series. Analysis and Forecasting. Narosa Publishing House Pvt. Ltd., New Delhi.

[21]. Cryer, J. D. 1986. Time Series Analysis. PWS Publishing, USA.

[22]. Rachana, W., Suvarna, M., Sonal, G. 2010. Use of ARIMA models for forecasting pigeon pea production in India. Int. Rev. Bus. Financ., 2(1): 97-107.